

Robust Laparoscopic Instruments Tracking Using Colored Strips

Virginia Mamone¹ (✉), Rosanna Maria Viglialoro¹, Fabrizio Cutolo¹, Filippo Cavallo², Vincenzo Ferrari¹

¹EndoCAS Center for computer-assisted surgery, Pisa, Italy

virginiam989@gmail.com

{rosanna.viglialoro, fabrizio.cutolo,
vincenzo.ferrari}@endocas.org

²Sant'Anna School of Advanced Studies, Pisa Italy

filippo.cavallo@santannapisa.it

Abstract. To assist surgeons in the acquisition of the required skills for the proper execution of the laparoscopic procedure, surgical simulators are used. During training with simulators it is useful to provide a surgical performance quantitative evaluation. Recent research works showed that such evaluation can be obtained by tracking the laparoscopic instruments, using only the images provided by the laparoscope and without hindering the surgical scene. In this work the state of the art method is improved so that a robust tracking can run even with the noisy background provided by realistic simulators. The method was validated by comparison with the tracking of a “chess-board” pattern and following tests were performed to check the robustness of the developed algorithm. Despite the noisy environment, the implemented method was found to be able to track the tip of the surgical instrument with a good accuracy compared to the other studies in the literature.

Keywords: optical tracking · single camera · laparoscopic training · surgical simulation · surgical performance evaluation.

1 INTRODUCTION

In recent decades, laparoscopic surgery proved to be an effective alternative to open surgery techniques for the treatment of various abdominal diseases. However, the laparoscopic technique introduces new issues for surgeons, due the using of the laparoscope and elongated instruments passing through the access ports, as the availability of the sole endoscopic camera mediated view, the lack of tactile feedback and the limited force feedback, and the fulcrum effect on the instruments that imposes restrictions on the movements [1]. These technique-related issues impose on novices a complex and long learning curve. Surgical simulators can shorten the training period and reduce patient's risks in the early stages of the learning curve [2, 3]. Surgical simulators allow novices to improve their technical abilities as manual dexterity and

the hand-eye coordination, and, for some simulators, learn the main difficult steps up to a complete operation. Through the analysis of certain parameters, it is possible to obtain a surgical performance quantitative evaluation, which constitutes a feedback of the progress and can be used to determine the achieving of an appropriate level of competence (proficiency level). Novices who have reached the proficiency level are ready, while other ones can need additional training. A literature search showed that significant parameters for this purpose, as path length and jerk [4], can be extracted from the movements of the laparoscopic instruments. In case of virtual reality simulators [5-7] the kinematics of the instruments is offered by the simulator itself, while in case of physical simulators (for example box trainer or mannequins) or hybrid AR simulators [5, 8] laparoscopic instruments tracking is required.

Since the laparoscopic images are available both in physical and hybrid AR simulators, machine vision based techniques are a suitable choice.

In [9] the authors combine machine learning methods with optical flow information to track the 3-D pose of the end-effector of a surgical robot for laparoscopy. Through a completely marker-less tracking, they are able to determine the 6 degrees of freedom (DoF) of the instrument in respect to the camera but the achievable tracking accuracy of 4.09 [mm] might be not enough for our purposes. In [10], laparoscopic surgical instruments are equipped with a non-bulky marker consisting in a colored strip attached to the instrument shaft. Machine vision methods allow the estimation of the projected diameter of the surgical instrument onto the camera images; taking into account the cylindrical shape of the surgical instrument shaft, this information is then used to determine the position of the instrument and its rotation along pitch and yaw angles. Mean tracking accuracy in [10] is 2.5 mm “for working distances commonly found in laparoscopic training”.

Our new approach improves the technique described in [10] to work with a more realistic and complex environment as the one provided by the cholecystectomy simulator being developed by the EndoCAS center for computer assisted surgery [11-13].

This paper describes the state of the art algorithm [10] and our improvements, the trial setup and the results. Finally the method and its application are discussed.

2 MATERIALS AND METHODS

2.1 Experimental set-up and camera calibration

The tests were carried out using the laparoscopic simulator being developed by the EndoCAS center for computer-assisted surgery [11-13]. It is an augmented reality simulator for the training of the critical phases in the cholecystectomy procedure. The simulator is composed of an external frame that replicates an insufflated abdomen while inside it faithfully reproduces the internal organs involved in the intervention. To get a more realistic render, liver and gallbladder are manufactured starting from computed tomography images [11, 14]. A laparoscope is reproduced through mounting on the distal part of a cylindrical tube, a consumer camera with a 4-6 cm focal length and a 56° viewing angle (REF-37131-919 by SOMIKON). The images are acquired at a 1280 x 720 resolution.

Colored strip implementation.

Marker detection is based on a simple color segmentation by thresholding performed in the HSV (hue, saturation, value) color space because HSV allows robust segmentation of objects that undergo non-uniform levels of illumination intensity, shadows and shading [10, 15]. In particular, segmentation based on saturation and hue allows wider color range to be covered and therefore it is less dependent on the lighting conditions within the scene (mainly affecting the V channel). Furthermore, highly saturated colors are not present in our simulated surgical scene (Fig. 1). For these reasons, and as suggested in [16] and [17], we adopted fluorescent dyes for the colored markers (strips) so as to peak the saturation channel in their own region of interest: the S values of the colored strips show a higher dynamic with respect to the neighboring pixels than their corresponding H values. In this way, the segmentation based on thresholding becomes more selective: it can be performed with a high cutoff value in the S-channel and it yields good results also at low lighting conditions like ours.

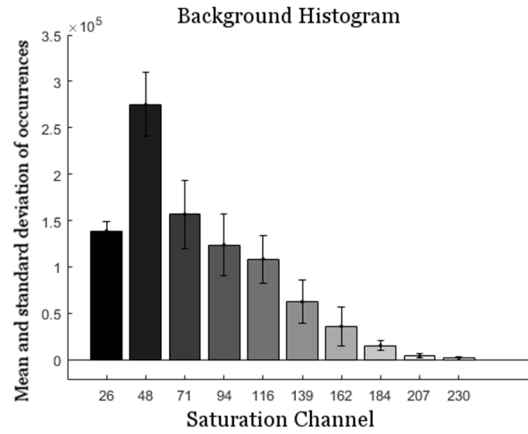


Fig. 1. Histogram of the Saturation channel (HSV-scale) of the images of the simulated surgical field: High saturated colors are infrequent.

Two laparoscopic grasping forceps, typically used during a cholecystectomy procedure, were modified by attaching to the distal part of each shaft a fluorescent colored strip. The colors were selected by analyzing the histogram of the hue channel of the operative field shown in figure 2. Red and, to a lesser extent, cyan prevail in the images of the operative field. Although the predominant color varies when different portions of the operative field are focused, the variance of the lower peaks is minimal. On the basis of these observations, a green marker and a blue marker were depicted with fluorescent dye respectively to the two shafts.

Camera calibration.

The camera is calibrated using the OpenCV libraries. The projective parameters of the camera and the distortion components due to lens non-linearity are estimated using the Zhang calibration procedure [18] and a “chessboard” pattern as calibration target.

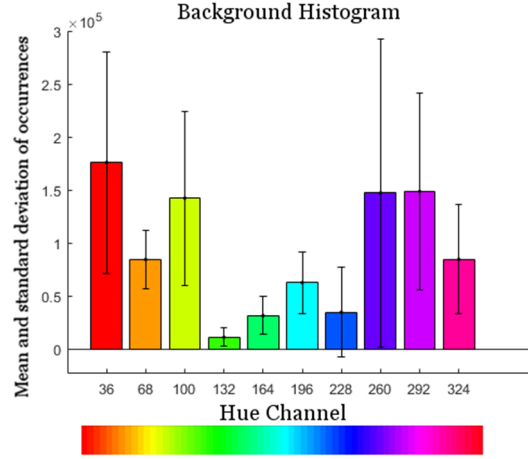


Fig. 2. Histogram of the Hue channel (HSV-scale) of the images of the simulated surgical field (background)

2.2 Image processing

The non-linear distortions onto the camera frames are compensated by applying an un-distortion function on the base of the calibration parameters.

The undistorted image is then processed by the two threshold-based segmentation procedures to detect the blue and green marker regions respectively. Segmentation returns a binary image but does not yield any topological information. Thereafter, a labeling technique is used to uniquely label the subsets of connected components. Among the identified regions, the one with the greatest area is associated to the marker mask, while the others are discarded. This will eliminate segmented areas not belonging to the marker. These areas may arise from the reflections of the marker on the metal surface of surgical instruments or on the silicon surface that makes up the internal organs (Fig. 3).

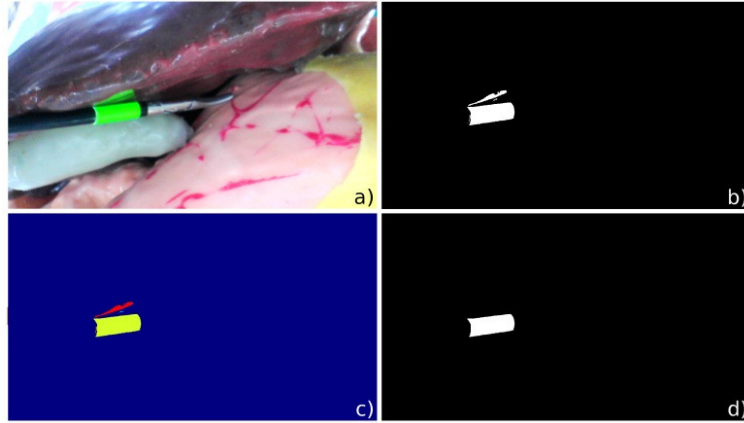


Fig. 3. a) An image captured by the simulator laparoscope. b) Image segmentation is altered by the reflection. c) Labeling algorithm discerns the topologically separate areas. d) Marker mask not affected by reflections in the starting image.

At this stage in [10], the authors apply the Hough transform to detect the straight lines associated to the profiles of the entire instrument. In the first video frame, they apply the Hough transform to the whole image and then in a sub-window taking into account the output of a Kalman filter to speed up the process. This step is not optimal for our set-up due to the inhomogeneous background. As showed in Fig. 4 (up) false straight lines not belonging to the instrument profile may be wrongly recognized by the Hough algorithm.

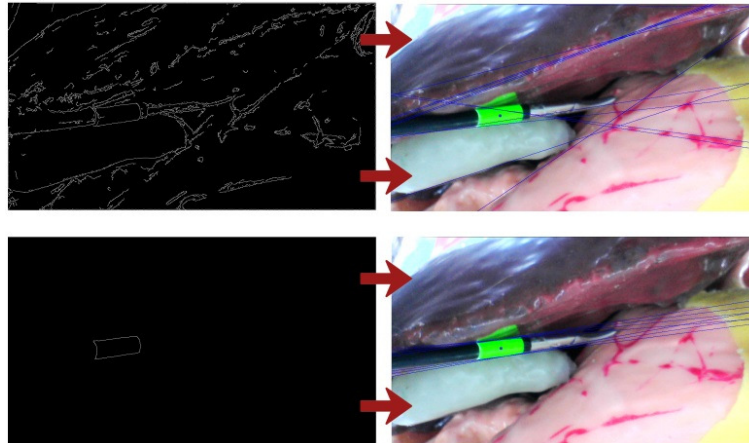


Fig. 4. (Up) application of the Canny detector to the whole image and corresponding identification of straight lines through Hough transform. (Down) application of the Canny detector to the segmented marker region and the corresponding identification of straight lines through Hough transform

In laparoscopic operations, the field of view (FOV) is usually very strict in order to magnify the view itself (Fig. 5). Therefore, the estimation of the lines passing through the profiles of markers is a good approximation of the profiles of the entire laparoscopic instrument. For this reason, we determine the straight lines applying the Hough transform only to the segmented regions corresponding to the markers. In this way, we drastically reduce the computational cost and, at the same time, we find only straight lines corresponding to the laparoscopic instrument (and not corresponding to the background as in the original version).

In the original version, the two straight lines with higher value in the Hough transform accumulator are identified as the two marker profiles. Nevertheless, it may happen that the two lines with the highest Hough transform are being associated to the same instrument profile (Fig. 5). When this happens, instrument tracking, based on instrument profiles knowledge (as described in the next paragraph), cannot be rightly performed.

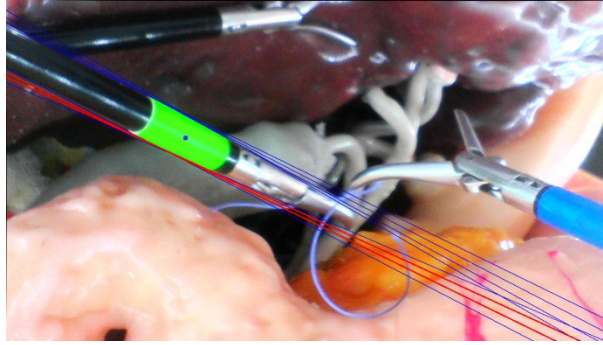


Fig. 5. Wrong identification of the instrument profiles selecting the 2 lines with higher Hough transform value: both red lines correspond to a single profile.

For this reason, we improved the method to prevent wrong labeling of the marker profiles (Fig. 6). At first, it is assumed that the straight line with the highest accumulator value corresponds to one of the two instrument profiles (*profile 1*). Then, the algorithm determines the *perpendicular line* to the identified *profile 1* and passing through the *barycenter* of the region corresponding to the marker (detected after segmentation). The *points of intersection* of the *perpendicular line* with all the straight lines returned by the Hough transform are identified. Finally, the *points of intersection* are divided into two groups via K-mean clustering. The two clusters of points appear on both profiles of the surgical instrument: they identify the straight lines associated to the opposite profiles 1 and 2. In this way, it is possible to identify the straight lines associated to *profile 2* as those contained in the second cluster of points (which does not contain *profile 1*). The straight line relative to *profile 2* is then chosen from this second cluster selecting the straight line with the highest value of the Hough transform.

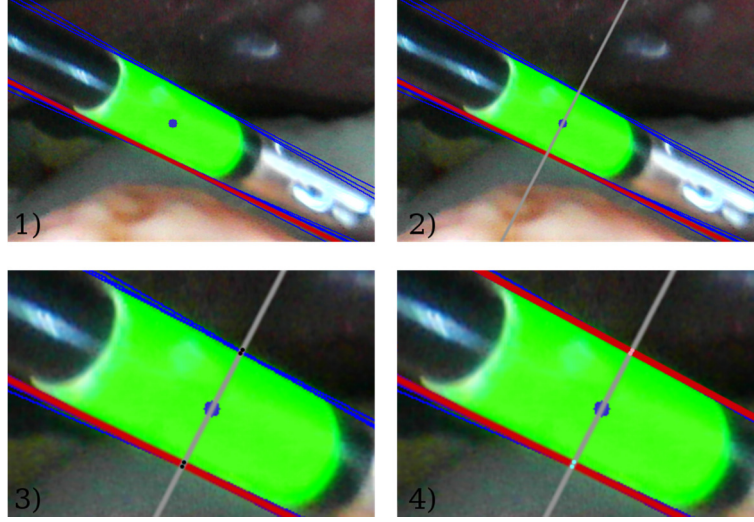


Fig. 6. Opposite profiles identification improvement: 1) Identification of the *profile 1*; 2) detection of the *perpendicular line* passing through the *barycenter* of the segmented region; 3) Intersection of the *perpendicular line* with all the straight lines returned by the Hough transform; 4) Profile 2 is identified via K-mean clustering.

2.3 Tracking of the tip of the surgical instrument

The apparent diameters of the surgical instrument are used to estimate the marker depth. The *main axis* of the instrument on the image is calculated as the median line between *profile 1* and *profile 2*. The points of intersection of the *main axis* with the segmented region of the marker determine points *A* and *B* (Fig. 7). Then it is possible to determine the lines perpendiculars to the main axis and passing through *A* and *B*. These lines intersect the instrument profiles respectively in A_1, A_2 and B_1, B_2 . Segments $\overline{A_1A_2}$ and $\overline{B_1B_2}$ approximatively correspond to the projection of the marker diameter in correspondence of *A* and *B*, and can be used to estimate the depth of their corresponding position in the 3-D space, as expressed in [10]:

$$Z_{A'} \approx d \frac{f_x f_y}{\sqrt{[(u_{A_1} - u_{A_2})f_y]^2 + [(v_{A_1} - v_{A_2})f_x]^2}}, \quad (1)$$

where d is the actual diameter of the laparoscopic instrument, f_x and f_y , are the horizontal and vertical focal lengths respectively, while $u_{A_1}, v_{A_1}, u_{A_2}$ and v_{A_2} are the coordinates on the image corresponding to A_1 and A_2 : $A_1 = (u_{A_1}, v_{A_1})$, $A_2 = (u_{A_2}, v_{A_2})$. After the estimation of $Z_{A'}$ it is possible to determine:

$$X_{A'} = (u_A - c_x) \frac{Z_{A'}}{f_x}, \quad (2)$$

$$Y_{A'} = (v_A - c_y) \frac{Z_{A'}}{f_y}, \quad (3)$$

where u_A , and v_A are the coordinates on the image corresponding to A , while c_x and c_y are the horizontal and vertical displacements of the optical axis from the center of coordinates.

Using the same equations it is possible to determine $X_{B'}$, $Y_{B'}$, and $Z_{B'}$.

Points A' and B' identify a vector in the 3D space. Considering the cylindrical shape of the laparoscopic instrument, its axis corresponds to the straight line identified by vector $\overline{A'B'}$ and the instrument tip lies on this line. Knowing the actual distance between the instrument tip P' and the marker, its spatial coordinates are trivially derived (Fig. 7).

We can perform a sanity check after the estimation of the tip position P' on each image frame taking into account current and previous tip positions. The sanity check is performed discharging tip positions estimations not compatible with the physical dimensions of the simulator or in case of too fast movements.

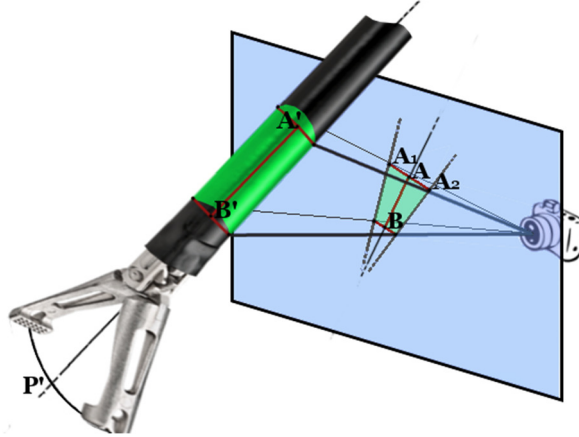


Fig. 7. Perspective projection of the points identified for estimating the position of the instrument tip (point P') and corresponding points in 3-D space

Tracking correction at the image edges.

In general, the whole marker should be visible to perform surgical instrument tracking. However, the implemented method works as long as the Hough transform is able to identify the profiles of the instrument. Applying some precautions, the tracking of the marker can proceed even when the latter is at the borders of the camera FOV, and it is only partially visible in the scene. In these conditions, the straight line at the image border may be wrongly identified as *profile 1* of the instrument (Fig. 8a). The lines whose Hough parameters fall within the ranges of the image edges are therefore excluded in advance. In this way the colored marker can be localized even if it is only partially displayed in the scene (Fig. 8b).

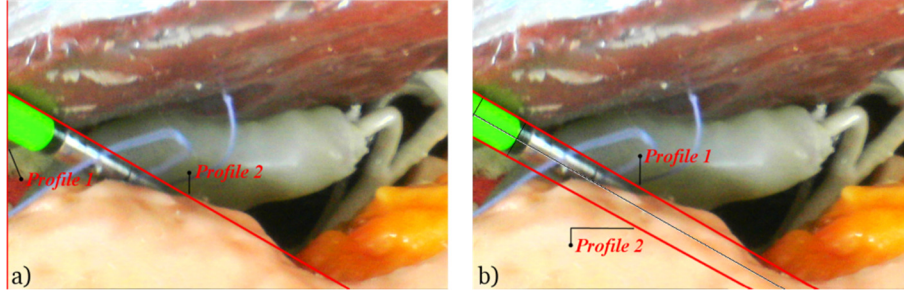


Fig. 8. Particular case with the marker partially contained in camera FOV. a) Wrong detection of the profiles of the laparoscopic instrument on the image edge. b) Proper identification of the profiles as a result of the exclusion of straight lines generated by the image edges

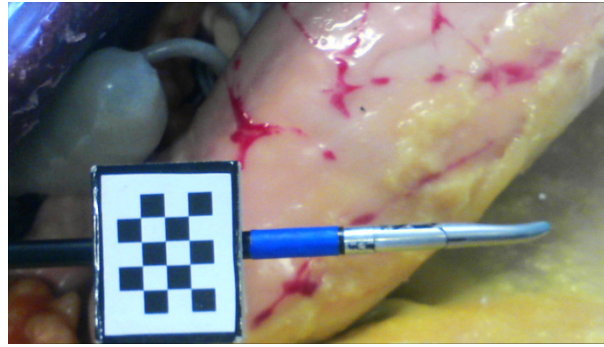


Fig. 9. Pattern used for obtaining the reference estimates

2.4 Accuracy evaluation

The following section describes the validation tests performed in the realistic environment provided by the cholecystectomy simulator. Our solution has been validated through using a “chessboard” pattern, commonly used in AR applications, as gold standard. As shown in Figure 9, the pattern was secured on a rigid support and attached to the laparoscopic instrument. Before proceeding with the test, the position of the laparoscopic instrument tip was mechanically calibrated (with a caliper) to both the marker and the grid reference. This configuration allows a direct comparison between the results obtained from the two tracking methods. The validation experiment was performed by randomly moving inside the cholecystectomy simulator and by simultaneously tracking the colored marker and the “chessboard” pattern. A total of 100 images were acquired and processed by the algorithm.

2.5 Robustness evaluation

Experimental tests were conducted to analyze the robustness and suitability of the method for the evaluation of surgeons in the laparoscopic cholecystectomy interven-

tion. The test involved two surgeons while performing the ligature of the cystic duct, a fundamental step of this surgical procedure. The test required the use of two laparoscopic grasping forceps, which were modified through the application of the colored markers. Figure 10 shows few frames captured during the test and highlighting the fundamental steps in the execution of the surgical knot. The trajectory of the tip of the surgical instrument during the simulation was processed to extract some parameters for the surgical performance evaluation. Path length D , normalized Jerk J , as well as task execution time T and time spent outside the camera FOV are calculated. The Path length D is the total length of the path taken by the tip of the surgical instrument. Chmarra et al. [19] have shown that this value is able to discern surgeons groups with different experience and is calculated as:

$$D = \int_0^T \sqrt{\left(\frac{dX}{dt}\right)^2 + \left(\frac{dY}{dt}\right)^2 + \left(\frac{dZ}{dt}\right)^2} dt . \quad (4)$$

The Jerk J characterizes the smoothness of the movement and its normalized form [4] is obtained as:

$$J = \sqrt{\frac{T^5}{2D^2} \int_0^T \left[\left(\frac{d^3X}{dt^3}\right)^2 + \left(\frac{d^3Y}{dt^3}\right)^2 + \left(\frac{d^3Z}{dt^3}\right)^2 \right] dt} . \quad (5)$$

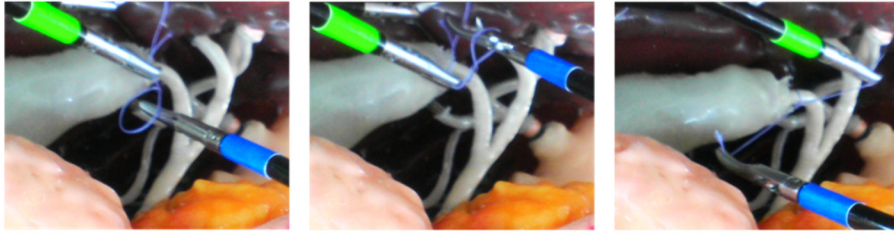


Fig. 10. Pictures taken during the experimental tests: The three sequential images illustrate the basic steps in the surgical knot execution.

3 RESULTS

3.1 Accuracy

The accuracy of the tracking method was evaluated by analyzing the deviation of the tip coordinates obtained through our tracking method from the coordinates obtained through the gold standard method. Figure 11 shows points corresponding to the laparoscopic instrument tip. Pink points result from the "chessboard" pattern tracking whereas blue points result from the colored marker tracking method. Table 1 lists the average error and the standard deviation obtained via the new tracking method. Much of the data variability can be attributed to the Z coordinate error. Although the average

error along the three directions is comparable, the standard deviation along Z is greater. This is due to the different way by which the depth is estimated: Z coordinate is estimated from the apparent diameter at the ends of the marker, the others are directly derived by the position of the 3D point in the image. Depth resolve through projective images is always an issue.

Table 1. The mean error and the standard deviation of each translational degree of freedom

	X	Y	Z
Mean	0.55 [mm]	0.24 [mm]	1.77 [mm]
Standard deviation	4.07 [mm]	1.82 [mm]	16.93 [mm]

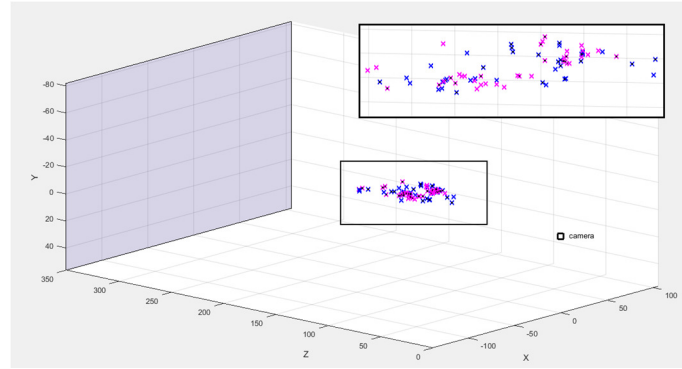


Fig. 11. Method quantitative validation results: Pink dots correspond to the instrument tip position obtained through the "chessboard" pattern. Blue points show the instrument tip position obtained through the colored marker.

3.2 Robustness

Figure 12 shows an image processed by the algorithm. The arrows on the image plane correspond to vectors \overrightarrow{AB} of each laparoscopic instrument as a result of the image processing. The figure highlights the projective geometry related to the 3-D localization of points A' and B' . The black straight lines represent the projection rays of points A and B . The Z coordinate estimate allows the localization of A' and B' in 3-D space therefore identifying $\overrightarrow{A'B'}$ vectors. The vectors are shown by the green and the blue arrows according to the color of the corresponding marker. The laparoscopic instruments tips, represented by points P_L and P_R in figure 12, are located along the direction of these vectors.

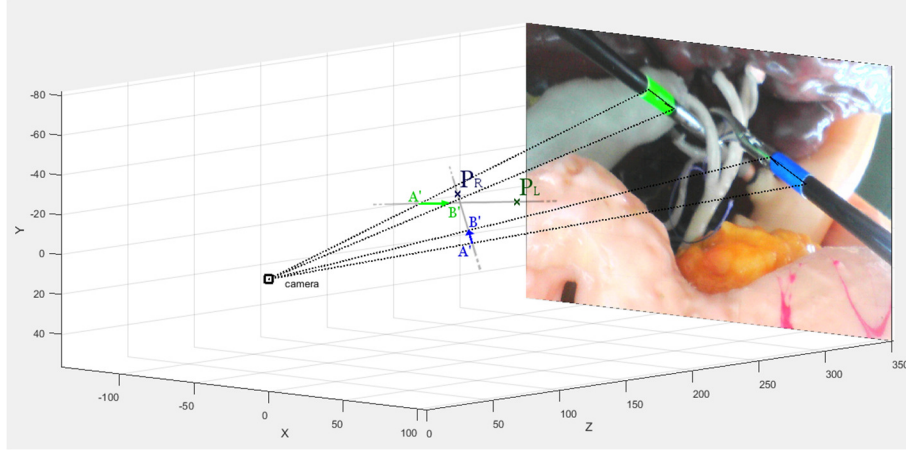


Fig. 12. Reconstruction process of the 3-D position starting from the identification of the colored markers in the image

To verify the reliability of the algorithm, the frames captured during the experimental tests were divided into three groups. The first group includes the images where the marker is completely in the scene without any parts outside the camera FOV. The frames where the marker is not properly localized can be considered as *false negatives*. In the second group, the images where the marker is completely absent from the scene are collected. In this case, the frames in which the marker is localized correspond to *false positives*. Finally, the third group contains frames in which the marker is only partially displayed (because a part of it falls outside of the camera FOV or it is hidden by the other instrument or by an anatomical part). In this group, the images in which the marker was correctly located or was identified as external to the scene are assessed. Table 2 summarizes the results obtained for each colored marker.

False negatives are mostly limited to cases of motion blur, while thanks to the use of fluorescent paint for the colored markers, *false positives* has not occurred. Regarding the last group of frames, the visible part of the marker is not always enough to provide the data needed to a proper localization and so the performance is very bad in respect to the previous two groups, but, in any case, marker visibility should be a must in case of marker based tracking.

The tracking method has proven robust and through the processing of the tip trajectory it was possible to obtain the parameters characterizing the surgical performance. The extracted parameters are shown in Table 3.

Table 2. Error statistics: *False negatives* correspond to cases in which the marker is wrongly localized although it is completely displayed in the image. *False positives* correspond to cases where the marker 3D position is located within the simulator though it does not appear within the image. *Wrong partial detection* refers to frames in which the marker, partially displayed in the image, has not been correctly located.

Experiment run	Surgeon 1		Surgeon 2	
	Green marker	Blue marker	Green marker	Blue marker
Error statistics				
False negative	0 %	0.18 %	0.96 %	0 %
False positive	0 %	0 %	0 %	0 %
Wrong partial detection	31.67 %	23.83 %	13.8 %	12.66 %

Table 3. Parameters for the evaluation of the surgical performance

Experiment run	Surgeon 1		Surgeon 2	
	Green marker	Blue marker	Green marker	Blue marker
Parameters				
Path length [mm]	3530	3790	1580	1360
Time outside FOV [s]	0.5	8.5	0.5	1
Normalized Jerk	$9.36 \cdot 10^5$	$1.19 \cdot 10^6$	$8.56 \cdot 10^6$	$6.50 \cdot 10^5$
Time [s]	106		78	

4 CONCLUSIONS

In this work some fundamental improvements in respect to a previous work for the tracking of laparoscopic surgical tools using colored strips are described. The implemented method is consistent with the results obtained from [10] although the latter uses a simulator that provides a background less noisy that simplifies the tracking. Our improvements can be summarized in: 1) identification of the straight lines in the segmented region of the marker and not on the entire image; 2) use of the K-mean clustering to rightly identify the two profiles; 3) Tracking correction at the image edges.

Thanks to our improvements, the method allows performing the tracking of surgical instruments even in noisy and realistic environments such like the one provided by the cholecystectomy simulator developed at the EndoCAS center.

Acknowledgement. This research work was supported by VALVETECH project, FAS fund – Tuscany Region (Realization of a newly developed polymeric aortic valve, implantable through robotic platform with minimally invasive surgical techniques) and SThARS project, grant “Ricerca finalizzata e Giovani ricercatori 2011-2012” Young Researchers - Italian Ministry of Health (Surgical Training in identification and isolation of deformable tubular structures with hybrid Augmented Reality Simulation).

1. Freschi, C., et al., *Technical review of the da Vinci surgical telemanipulator*. International Journal of Medical Robotics and Computer Assisted Surgery, 2013. **9**(4): p. 396-406.
2. Champion, H.R. and A.G. Gallagher, *Surgical simulation—a 'good idea whose time has come'*. British Journal of Surgery, 2003. **90**(7): p. 767-768.
3. Kneebone, R., *Simulation in surgical training: educational issues and practical implications*. Med Educ, 2003. **37**(3): p. 267-77.
4. Cavallo, F., *Surgical Gesture Analysis: Systems and methods for biomechanical modelling of surgeon in minimally invasive surgery*, in *Scuola Superiore Sant'Anna*. 2007, Pisa.
5. Lahanas, V., E. Georgiou, and C. Loukas, *Surgical Simulation Training Systems: Box Trainers, Virtual Reality and Augmented Reality Simulators*, in *Int J Adv Robot Automn*. 2016. p. 9.
6. Moglia, A., et al., *Patient specific surgical simulator for the evaluation of the movability of bimanual robotic arms*. Stud Health Technol Inform, 2011. **163**: p. 379-85.
7. Turini, G., et al., *Patient-specific surgical simulator for the pre-operative planning of single-incision laparoscopic surgery with bimanual robots*. Comput Aided Surg, 2012. **17**(3): p. 103-12.
8. Ferrari, V., et al., *Augmented reality visualization of deformable tubular structures for surgical simulation*. The International Journal of Medical Robotics and Computer Assisted Surgery, 2015.
9. Allan, M., et al., *Image Based Surgical Instrument Pose Estimation with Multi-class Labelling and Optical Flow*, in *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I*, N. Navab, et al., Editors. 2015, Springer International Publishing: Cham. p. 331-338.
10. Loukas, C., V. Lahanas, and E. Georgiou, *An integrated approach to endoscopic instrument tracking for augmented reality applications in surgical simulation training*. Int J Med Robot, 2013. **9**(4): p. e34-51.
11. Viglialoro, R.M., et al., *Augmented Reality Simulator for Laparoscopic Cholecystectomy Training*, in *Augmented and Virtual Reality: First International Conference, AVR 2014, Lecce, Italy, September 17-20, 2014, Revised Selected Papers*, L.T. De Paolis and A. Mongelli, Editors. 2014, Springer International Publishing: Cham. p. 428-433.
12. Viglialoro, R.M., et al. *AR Visualization of "Synthetic Calot's Triangle" for Training in Cholecystectomy*. in *12th IASTED Int. Conf. on Biomedical Engineering*. 2016. Austria, Innsbruck: BioMed
13. Viglialoro, R.M., et al. *A physical patient specific simulator for cholecystectomy training*. in *Computer Assisted Radiology and Surgery (CARS)*. 2012.
14. Condino, S., et al., *How to build patient-specific synthetic abdominal anatomies. An innovative approach from physical toward hybrid surgical simulators*. Int J Med Robot, 2011. **7**(2): p. 202-13.

15. Kyriakoulis, N. and A. Gasteratos. *Light-invariant 3D object's pose estimation using color distance transform*. in *Imaging Systems and Techniques (IST), IEEE International Conference*. 2010
16. Diotte, B., et al., *Multi-modal intra-operative navigation during distal locking of intramedullary nails*. IEEE Trans Med Imaging, 2015. **34**(2): p. 487-95.
17. Cutolo, F., et al., *Robust and Accurate Algorithm for Wearable Stereoscopic Augmented Reality with Three Indistinguishable Markers.*, in *Electronics*. 2016. p. 3:59.
18. Zhang, Z., *A Flexible New Technique for Camera Calibration*. IEEE Trans. Pattern Anal. Mach. Intell., 2000. **22**(11): p. 1330-1334.
19. Chmarra, M.K., et al., *Retracting and seeking movements during laparoscopic goal-oriented movements. Is the shortest path length optimal?* Surg Endosc, 2008. **22**(4): p. 943-9.